

Methodological and conceptual challenges of evaluating the impact of development interventions

Citation for published version (APA):

Vaessen, J. (2010). *Methodological and conceptual challenges of evaluating the impact of development interventions*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20100916jv>

Document status and date:

Published: 01/01/2010

DOI:

[10.26481/dis.20100916jv](https://doi.org/10.26481/dis.20100916jv)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 05 May. 2023

SAMENVATTING

Inleiding

In de afgelopen tien jaar hebben verschillende ontwikkelingen in de internationale gemeenschap geleid tot een toenemende aandacht voor de effecten van de ontwikkelingsinterventies. Zowel binnen de internationale gemeenschap van ontwikkelingssamenwerking als daarbuiten heerste het gevoel dat de resultaten van ontwikkelingsinterventies tegenvielen in vergelijking tot de verwachtingen. Bovenal was het gevoel dat de bewijsvoering omtrent de effecten van ontwikkelingsinterventies zeer zwak was. Als gevolg hiervan zien we in de tweede helft van het eerste decennium van de 21^e eeuw een sterke toename in de aandacht en financiering voor impactevaluatie. Impactevaluatie kan in brede zin gedefinieerd worden als het geheel van evaluatieve praktijken dat zich richt op het vaststellen van beoogde en niet-beoogde resultaten van beleidsinterventies.

Een bijzondere vorm van impactevaluatie die veel aandacht krijgt betreft zogenaamde gerandomiseerde experimenten. Deze vorm van effectonderzoek wordt door velen beschouwd als de meest rigoureuze techniek om de effecten van beleidsinterventies vast te stellen. De discussie rond de toepassing van gerandomiseerde experimenten in ontwikkelingssamenwerking heeft in de afgelopen paar jaren geleid tot een grote controverse tussen de zogenaamde ‘randomistas’, voorstanders van gerandomiseerde experimenten, en tegenstanders. Een van de basisprobleemstellingen van dit proefschrift is dat deze controverse een averechts effect heeft op het bereiken van een doelstelling waar iedereen het wel over eens is: het accumuleren van kennis omtrent wat werkt en wat niet, en onder welke omstandigheden.

Hoofdstuk 1 beschrijft het overkoepelende kader van het proefschrift. In dit hoofdstuk worden drie essentiële (categorieën van) uitdagingen in impactevaluatie beschreven. Deze zijn achtereenvolgens: afbakening, attributie versus verklaring, en praktische uitdagingen in de opzet en uitvoering van impactevaluaties. Dit kader is ook de kapstok waaraan alle hoofdstukken van dit proefschrift aan kunnen worden opgehangen.

In de bespreking van de drie uitdagingen wordt voortdurend de koppeling gemaakt met gerandomiseerde experimenten. Dit proefschrift toont aan dat gerandomiseerde experimenten niet geschikt zijn om op al deze uitdagingen een bevredigend antwoord te bieden. Tegelijkertijd wordt ook geargumenteed dat gerandomiseerde experimenten zeker een prominente plaats verdienen binnen het gamma van methodologieën van effectonderzoek. Het doel van dit proefschrift is het illustreren en bekrachtigen van de volgende stellingen:

- De vraag van wat werkt en waarom binnen het beleidsveld van ontwikkelingssamenwerking kan niet bevredigend worden beantwoord op basis van een enkele onderzoeksopzet (i.e. gerandomiseerde experimenten);
- Er zijn een aantal factoren die de toepassing van gerandomiseerde experimenten beperken;

- De kwaliteit van impactevaluaties kan worden verbeterd door deze te kaderen in een beleidstheoretische benadering.

Hoewel de titel van dit proefschrift een algemene beschouwing omtrent bovenstaande punten impliceert, zijn omwille van praktische redenen een aantal keuzes gemaakt omtrent de interventies, beleidsvelden en contexten die worden behandeld in dit boek. Een drietal elementen kunnen worden onderscheiden die een verdere afbakening van dit proefschrift vormgeven. Ten eerste zijn de hoofdstukken 4, 5 en 6 gebaseerd op empirische analyses van beleidsinterventies op het gebied van (duurzame) landbouw gericht op kleine en middelgrote boeren. In hoofdstukken 2 en 3 wordt ook gesproken over interventies op het gebied van landbouw en milieu. Ten tweede is er ook een geografische focus in de empirische analyses. Hoofdstuk 4 is gebaseerd op veldwerk in Nicaragua (en Costa Rica) en hoofdstukken 5 en 6 komen voort uit veldwerk in Guatemala. Het derde element betreft de institutionele context van beleidsevaluatie. Hoofdstukken 2 en 4 gaan over interventies gefinancierd door de *Global Environment Facility*, een globaal fonds dat projecten steunt in ontwikkelingslanden op het gebied van biodiversiteit, klimaatsverandering en andere milieudoelstellingen.

Uitdaging 1: afbakening

Hetgeen wel of niet wordt bestudeerd binnen het kader van een impactevaluatie kan heel erg verschillen naar gelang de aard van het evaluandum, de aard van de effecten die kunnen optreden, als ook verdere keuzes die worden gemaakt omtrent de aspecten die meegenomen moeten worden. In hoofdstuk 1 komen de volgende aspecten aan de orde: het omgaan met uiteenlopende waarden en prioriteiten van de verschillende groepen van betrokkenen bij een evaluatie, de vraag van ‘effecten van wat?’, en de vraag van ‘effecten op wat?’ Deze aspecten komen in verschillende hoofdstukken aan de orde.

Hoofdstuk 2 behandelt het thema van methodologische uitdagingen in het evalueren van de effecten van een portfolio van interventies op het gebied van biodiversiteit. Gegeven de heterogeniteit binnen de portfolio en budgettaire restricties is het belangrijk om te reflecteren over het spanningsveld tussen ‘breedte’ en ‘diepte’ van de analyse. Ter ondersteuning worden de volgende twee vragen gesteld: de ‘effecten van wat’ en de ‘effecten op wat?’ Verschillende aspecten van afbakening komen aan de orde. Een belangrijk aspect is niveau van analyse. In de keuze van niveau van analyse kan men een onderscheid maken tussen administratieve categorieën zoals projecten of programma’s en analytische categorieën zoals strategische prioriteiten en beleidsinstrumenten die teruggevonden kunnen worden in meerdere interventies. Niveau van analyse heeft ook betrekking op de effecten. Het maakt nogal een verschil in termen van onderzoeksopzet of men kijkt naar veranderingen op institutioneel en individueel niveau of op het niveau van milieudoelstellingen. De causaliteit tussen beleidsinterventies en dit laatste niveau is veel indirecter en diffuser dan in het geval van individueel en institutioneel gedrag. Hoofdstuk 2 be-

handelt ook de beleidstheoretische evaluatiebenadering als kader voor afbakening en analyse op verschillende niveaus.

Hoofdstuk 3 bespreekt in detail de beleidstheoretische evaluatiebenadering. Met betrekking tot het thema van afbakening gaat het hoofdstuk in op een specifiek kader voor het deconstrueren van veranderingsprocessen naar verschillende typen van mechanismen. Deze mechanismen betreffen achtereenvolgens macro-micro-, micro-micro-, en micro-macro-veranderingsprocessen. Het hoofdstuk laat zien hoe een dergelijke deconstructie nuttig kan zijn in het structureren van de verzameling en analyse van data binnen een interpretatiekader van hoe veranderingsprocessen kunnen verlopen.

Hoofdstuk 4 betreft een analyse van een gerandomiseerd experiment in het kader van een ontwikkelingsinterventie in Nicaragua, Costa Rica en Colombia. Op systematische wijze wordt het potentieel van het experiment om antwoord te geven op vragen omtrent de effectiviteit van monetaire incentives en technische assistentie op landgebruiksveranderingen onderzocht. Gerandomiseerde experimenten zijn normaliter uitermate geschikt om directe effecten ('outcomes') vast te stellen. In dit hoofdstuk wordt ook bekeken in hoeverre het experiment iets kan zeggen over impact in brede zin zoals gedefinieerd door de *OESO*. In het kader van deze oefening worden verschillende beperkingen vastgesteld en concrete suggesties voor additionele vormen van analyse worden besproken. Voorbeelden van effecten buiten het bereik van gerandomiseerde experimenten zijn zogenaamde 'verplaatsingseffecten' (effecten op landgebruik elders buiten de zone van interventie maar wel als gevolg van de interventie) en de duurzaamheid van landgebruiksveranderingen en dieperliggende effecten op milieudoelstellingen.

Hoofdstuk 6 presenteert een methodologisch kader voor impactevaluatie vanuit het oogpunt van ex ante effectonderzoek van beleidsalternatieven. Een belangrijke veronderstelling in dit onderzoek is dat enkel de bewijsvoering rond potentiële effecten onvoldoende is als basis voor toekomstig beleid. Er moet ook rekening worden gehouden met de prioriteiten van verschillende groepen van betrokkenen. Het methodologische kader dat wordt uitgewerkt in dit hoofdstuk maakt het mogelijk om beleidsalternatieven met elkaar te vergelijken op basis van potentiële effecten, daarbij rekening houdend met verschillen in het relatieve belang van effecten voor verschillende groepen van betrokkenen.

Uitdaging 2 : attributie versus verklaring

Gerandomiseerde experimenten maken het mogelijk om de netto-effecten van een interventie vast te stellen met een hoge graad van interne validiteit. Echter, tegelijkertijd zijn er een aantal beperkingen in termen van het helpen verklaren van veranderingsprocessen. Een beleidstheoretische benadering kan een gerandomiseerd experiment op verschillende punten versterken. Hoofdstuk 1 behandelt een aantal aspecten onder dit thema, waaronder een korte bespreking van het attributievraagstuk en de rol van gerandomiseerde experimenten, de mogelijke wisselwerking tussen interne en externe validiteit in effectonderzoek, het belang van beleids-

theorieën, en de ‘wet van de comparatieve voordelen’ van verschillende onderzoeksmethodologieën.

De belangrijkste doelstelling van hoofdstuk 4 is het bestuderen van het potentieel van een gerandomiseerd experiment in de vaststelling van bepaalde effecten van een concrete interventie in drie landen in Latijns-Amerika. De analyse laat zien hoe verschillende tekortkomingen dit analytisch potentieel in gevaar brengen. Het experiment bestaat uit verschillende facetten. Twee van de drie vergelijkingen tussen interventie- en controlegroepen zijn door de verschillende problemen in het experiment niet meer betrouwbaar. De derde optie, de vergelijking tussen groepen met verschillende modaliteiten van monetaire incentives, biedt daarentegen een betrouwbaar kader voor het analyseren van netto-effecten. Daarnaast laat dit hoofdstuk ook zien hoe het analytische potentieel van een gerandomiseerd experiment versterkt kan worden met additionele methoden en de toepassing van een beleidstheoretische benadering.

Hoofdstukken 3 en 4 laten beiden op hun eigen wijze zien hoe een beleidstheoretische benadering in evaluatie de interne, externe en constructvaliditeit van bevindingen in effectonderzoek kan versterken. In hoofdstuk 3 wordt ingegaan op de vragen van wat een beleidstheorie precies is, hoe deze kan worden gereconstrueerd en hoe deze kan worden toegepast in evaluatie-onderzoek. Het hoofdstuk bespreekt ook hoe een specifieke versie van een beleidstheoretisch kader, gebaseerd op Coleman's *Theory of Social Action*, van nut kan zijn om veranderingsprocessen beter kunnen begrijpen. Dit kader biedt niet alleen een bruikbare structuur voor de verzameling en analyse van data maar ook een kader van interpretatie ingebed in een theorie van individueel en collectief gedrag.

Hoofdstuk 5 presenteert een impactevaluatie gebaseerd op kwantitatieve en kwalitatieve methoden en ingebed in een theoretisch kader. Het is een voorbeeld van een zogenaamde *shoestring approach*; de evaluatie is gericht op het genereren van maximaal betrouwbare gegevens en bevindingen binnen een context van beperkingen in budget, tijd en data.

Het principe dat verschillende methoden comparatieve voordelen hebben bij bepaalde aspecten of vragen binnen een impactevaluatie wordt ook geïllustreerd in hoofdstuk 6. In het geval van gerandomiseerde experimenten is de zogenaamde ‘counterfactual’ de situatie waarin er geen interventie plaatsvindt. Echter, in de praktijk zijn beleidsmakers geïnteresseerd in andere ‘counterfactuals’, namelijk een andere interventie (in plaats van geen interventie). Het methodologische kader dat wordt ontwikkeld in dit hoofdstuk bestaat uit drie bouwstenen: multicriteria-analyse, een inventaris van beleidsprioriteiten onder betrokkenen, en een beleidstheorie rond veranderingsprocessen. Door deze elementen te combineren kunnen beleidsalternatieven systematisch met elkaar worden vergeleken, wat uiteindelijk leidt tot een rangorde van alternatieven. De beleidstheorie speelt hierbij een cruciale rol in de ondersteuning van de ex ante evaluatie van beleidsalternatieven.

Uitdaging 3: praktische uitdagingen in de opzet en uitvoering van impactevaluaties

Kwaliteitsvolle impactevaluatie impliceert goed onderzoek. Huidige debatten rond impactevaluatie benadrukken de rol en keuze van onderzoeksopzet, i.e. gerandomiseerde experimenten. Echter, het genereren van goed onderzoek, in de context van gerandomiseerde experimenten en daarbuiten, hangt vooral af van de mate waarin een aantal praktische uitdagingen in de opzet en uitvoering van onderzoek op een bevredigende wijze kunnen worden aangepakt. In hoofdstuk 1 worden verschillende mogelijke tekortkomingen (of ‘dreigingen’) besproken die de interne validiteit van bevindingen van gerandomiseerde experimenten in gevaar kunnen brengen. Daarnaast worden een aantal concrete uitdagingen in de opzet en uitvoering van impactevaluaties besproken.

Hoofdstuk 4 analyseert de mogelijke tekortkomingen van een gerandomiseerd experiment die de interne validiteit van de bevindingen in gevaar kunnen brengen. De belangrijkste problemen die systematisch worden onderzocht zijn: bias in selectie, contaminatie, en verschillende onbedoelde gedragsveranderingen binnen de doelgroep. In het algemeen kan men stellen dat differentiatie tussen groepen van boeren binnen de experimentele opzet nogal wat onrust veroorzaakte binnen de doelgroep. Dit valt deels te wijten aan de communicatiestrategie in de beginfase van het project. De problemen bij de opzet en uitvoering van het gerandomiseerde experiment zijn in wezen terug te voeren naar een aantal basisoorzaken die betrekking hebben op de organisatie en kwaliteitscontrole van het experiment. Het project had een zeer goede professionele staf, met o.a. experts in ecologie, zoölogie en agromonie. De stafleden in het veld waren echter geen ervaren onderzoekers. Onderzoekers waren slechts in beperkte mate op bepaalde tijdstippen aanwezig. De staf in het veld had noch de kennis noch de incentives om een gerandomiseerd experiment tot een goed einde te brengen. In dit hoofdstuk worden verschillende praktische aanbevelingen verstrekt ter verbetering van de opzet en uitvoering van gerandomiseerde experimenten in soortgelijke projecten.

Hoofdstuk 5 betreft een impactevaluatie van een trainingsprogramma in organische landbouw in de hooglanden van Guatemala ten tijde (1998-2001) van een herstelperiode na een langdurige burgeroorlog. Gegeven structurele problemen van wantrouwen en taalbarrières (Spaans versus lokale Maya-talen) tussen onderzoekers (inclusief lokale onderzoekers) en boeren, is het onderzoek in sterke mate gebaseerd op het principe van triangulatie van bevindingen uit enquête-onderzoek, semi-gestructureerde interviews en veldbezoeken. Daarnaast is het onderzoek ingebed in een theoretisch kader gebaseerd op literatuur rond adoptie-gedrag en economisch gedrag van kleine boeren. Het onderzoek is gebaseerd op een simpele quasi-experimentele opzet bestaande uit drie gekoppelde steekproeven: participerende boeren ex ante en ex post en een controlegroep ex post. De reden waarom een controlegroep ex ante ontbreekt heeft te maken met de beperkte omvang van de steekproeven. Gegeven de hoge migratiecijfers (permanent en seizoensgebonden) en andere factoren zou het te moeilijk zijn geweest om in 2001 voldoende boeren uit de controlegroep te vinden die bereid zouden zijn om mee te werken aan het onder-

zoek. Dit probleem deed zich niet voor in de deelnemersgroep gegeven hun band met het project. Hoofdstuk 5 is een illustratief voorbeeld van hoe men bij de opzet en uitvoering van een impactevaluatie zou kunnen inspelen op verschillende praktische problemen teneinde zo betrouwbaar mogelijke resultaten te kunnen genereren in sub-optimale onderzoeksomstandigheden.

Tot slot

In het slothoofdstuk worden de initiële stellingen van dit proefschrift nogmaals kort onder de loep gehouden. Een aantal aspecten komen hierbij naar voren. Allereerst kan men stellen dat de vraag van ‘wel of niet gerandomiseerde experimenten’ teveel aandacht heeft gekregen in het debat. Tegelijkertijd moet ook niet het tegenovergestelde pad worden bewandeld door deze waardevolle methodologie aan de kant te schuiven. Het slothoofdstuk gaat verder in op de problematiek van beperkte toepasbaarheid van gerandomiseerde experimenten enerzijds en de kracht van een gerandomiseerd experiment anderszijds. In combinatie met een goed theoretisch kader en voortbouwend op de in vele contexten al bestaande interesse voor experimenteren met beleid (in velerlei vormen) verdienen gerandomiseerde experimenten een plaats in impactevaluatie. Een belangrijk indirect effect van het hele debat omtrent gerandomiseerde experimenten betreft de toegenomen aandacht voor kwaliteit in effectonderzoek. In het stijgende aantal impactevaluaties gebaseerd op een gamma van onderzoeksmethodologieën is er duidelijk aandacht voor kernuitdagingen zoals het attributie-vraagstuk en de externe validiteit van bevindingen. Andere belangrijke aspecten zoals de kwaliteit van dataverzameling worden echter onterecht nog steeds onvoldoende belicht.